

УДК 004.414.2

КОНТРОЛЬ КАЧЕСТВА ДАННЫХ С ПОМОЩЬЮ БЛОКЧЕЙН ТЕХНОЛОГИЙ

Токаев А.Д. (СПбГУ)

Научный руководитель – канд. физ.-мат. наук Корхов В.В.

(СПбГУ)

Введение. Проблема повышения качества и скорости обработки данных в распределенных системах на сегодняшний день неотложна. Низкое качество данных может привести к ошибочной информации и сбоям в информационных системах. Качество данных широко исследовалось в течение последних нескольких десятилетий [1] и к настоящему времени превратилось в профессиональную дисциплину. Растущий интерес к изучению параметров качества данных объясняется тем, что эффективность информационных систем зависит от качества используемых информационных ресурсов. Однако, существующие фундаментальные аспекты управления качеством данных не обеспечивают общего понимания основных знаний дисциплины. Исследования свидетельствуют о недостаточном контроле качества процесса отбора информации, что приводит к снижению качества данных. Современные технологии, такие как децентрализованные технологии реестра и искусственный интеллект, могут помочь решить проблему. Способность технологии блокчейн контролировать согласованность и целостность данных позволит эффективнее управлять качеством данных.

Качество данных (КД) - характеристика, которая показывает степень полезности данных [2]. Согласно стандартам ISO 9000: 2015, ключевыми критериями качества данных являются: релевантность, точность, целостность, полнота, последовательность, объем данных, своевременность. Общая оценка качества может быть получена посредством среднего балла атрибутов в рамках принятой модели качества для данной системы.

Основная часть. В докладе проводится анализ существующих методов контроля качества данных для задачи транспортировки жидкости с определенной температурой возгорания, перевозимой в специальных резервуарах, поддерживающих постоянные значения температуры и давления. Существующие блокчейн решения, позволяющие обеспечивать контроль качества данных в рассматриваемой задаче, реализуются на основе публичного блокчейна, что влечет за собой ряд минусов, такие как: низкая скорость транзакции, недостаточная конфиденциальность (данные, хранящиеся в блокчейне, находятся в общем доступе), высокие затраты на транзакции (публичные блокчейны взимают комиссию). Поэтому было принято решение реализации механизмов контроля качества данных на основе приватных блокчейнов. Оценка качества данных в блокчейне позволяет узлам выполнять проверки качества до того, как транзакции, несущие данные, будут зарегистрированы в блокчейне или использованы для вызова логики приложения блокчейна. Правила для определения качества данных могут быть реализованы в виде определенного класса смарт-контрактов.

Выводы. Реализация механизмов оценки качества данных в блокчейне расширяет возможности внедрения технологии блокчейна в межорганизационные информационные системы.

Список использованных источников:

1. Jayawardene, Vimuthki, Sadiq, Shazia, and Indulska, Marta An analysis of data quality dimensions // ITEE Technical Report. - 2015. - №2013-01 and 2015-027.
2. Degtyarev A., Bogdanov A., Shchegoleva N., Khvatov V. Data Quality in a Decentralized Environment // Lecture Notes in Computer Science. - 2020. - №12251. - С. 58-71.

Токаев А.Д. (автор)

Подпись

Корхов В.В. (научный руководитель)

Подпись